ROLE OF TRAFFIC OFFICERS IN TRANSPORTATION ASSET MONITORING

Tom Strain, Corresponding Author
Postgraduate Student
Department of Engineering Mathematics
University of Bristol, Ada Lovelace Building, University Walk, Bristol, United Kingdom, BS8 1TW
Email: tom.strain@bristol.ac.uk

R. Eddie Wilson
Professor of Intelligent Transport Systems
Department of Engineering Mathematics
University of Bristol, Ada Lovelace Building, University Walk, Bristol, United Kingdom, BS8 1TW
Email: re.wilson@bristol.ac.uk

Roger Littleworth
Independent Geospatial Consultant
Email: roger.littleworth@gmail.com

Word Count: 6145 words + 2 table(s) × 250 = 6645 words

Submission Date: July 22, 2021
ABSTRACT

In this paper, we investigate the feasibility of using highway traffic officers (TOs) for transportation asset management (TAM) alongside their primary role of incident response. Asset data, typically captured via highway surveys on an annual basis, is unsuitable for those assets whose condition might rapidly change, such as vegetation, street lights, guardrails, or drainage systems. Therefore, we consider as a proof-of-concept, whether data collected from dashboard cameras installed in TO vehicles might provide analysts with near real-time asset data across an entire highway network.

We consider a case study of a dedicated TO fleet deployed on the strategic road network (SRN) in England, UK, and develop a simulation based on publicly available data sets.

Within the simulation, TOs patrol under two distinct regimes and respond to dynamically generated incidents. The first regime aims to minimise the fleet’s incident response time, and the second aims to maximise the fleet’s coverage, with an aim to capture asset data across the entire highway network. Overall, our simulations show that the TOs deployed for TAM reduce the SRN junction-to-junction section inter-visit time by around 1 hour 45 minutes, whereas their incident response time is only increased by about 4 minutes. Moreover, 17% of SRN sections are not visited at all when the TOs prioritise fast incident response, which is reduced to 2% when the TOs prioritise the capture of asset data.

Keywords: Transportation Asset Management, Traffic Officers, Simulation
INTRODUCTION

Transportation asset management (TAM) is a key area of operation for a highway agency (1, 2) that ensures highway assets such as traffic signs and street lights are properly monitored and maintained. Effective TAM relies on up to date and accurate asset data to enable agency analysts to inspect each asset’s condition and thus design and implement suitable maintenance schedules.

Highway agencies usually deploy specialised vehicles, equipped with sensors such as a camera or LIDAR to perform remote surveys (3, 4), and thus an analyst may perform TAM from the safety of their office, rather than at the roadside. Further, a number of intelligent transport systems, such as the computer vision-based tools proposed by Strain et al. (5) and Golparvar-Fard et al. (6), have been developed to automatically analyse survey imagery and provide decision support and automation within the TAM process.

Highway surveys are usually performed on an annual basis, and are therefore not suitable for assets such as vegetation, drainage systems, street lights, and guardrails, whose condition might rapidly change. Therefore, in this paper, we consider as a proof-of-concept, whether Traffic Officers (TOs), whose primary role is to attend and manage traffic incidents, might also be used to collect asset data.

The apparent opportunity is those TO fleets that have a dedicated incident-management capability, separate from the police, and who are under the direct control of a highway agency. One such fleet is the Weginspecteurs (road stewards) deployed by Rijkswaterstaat (RWS), the national highway agency in the Netherlands (7). The RWS Weginspecteurs are trained by the Dutch police force in incident management and traffic control, thus enabling the police to focus on law enforcement. Similarly, Highways England (HE, the largest highway agency in the UK) Traffic Officers are a dedicated fleet deployed across the strategic road network (SRN) in England. The fleet has proven to be an effective capability since its introduction in 2004; contributing to a 12% reduction in incident-related congestion and freeing up 44% of police time (8).

As well as vehicle removal and traffic management equipment (e.g., tow ropes and cones) and communications systems (e.g., radio) (9), TO vehicles are often fitted with a dashboard camera (dash cam) (10), see Figure 1(c), and therefore represent a ubiquitous sensing capability on the highway. Monitoring the highway and its assets with such relatively low-quality but high-frequency imagery (compared against annual survey data) might compliment and improve upon the current TAM process by providing near real-time asset data to agency analysts.

In the work presented in this paper, we consider a case study of a dedicated TO fleet, who are likely to have relatively relaxed operational constraints compared with their police counterparts. To investigate the feasibility of the proposed TO-based TAM capability, a simulation is developed in which the TOs patrol along the highway and respond to dynamically generated incidents. Two distinct patrol regimes are considered, one that aims to minimise the fleet’s incident response time and one that aims to maximise the fleet’s coverage for asset management. By considering the incident response times and coverage achieved by each regime, we will determine the feasibility of deploying the TO fleet for TAM.

BACKGROUND AND LITERATURE REVIEW

Traffic Officer Fleets

Traffic Officers are usually formed from a police sub-unit. For example, the Autobahnpolizei patrol and manage incidents across the German autobahns (15), the Highway Patrol units (known as state troopers or state police in some states) are deployed on U.S. highways (16), and the Garda
FIGURE 1 Highways England (HE) traffic officers (a) and Weginspecteurs (f) respectively patrol along the strategic road network in England, UK (e) and the Dutch highway network. Each TO vehicle is equipped with a dashboard camera, shown in panel (c). Traffic officers are assigned to attend and manage traffic incidents (b) by regional control centre operators, shown in panel (d). Sources: (a), (b), and (d), Auto Express website (11); (c), online video frame (12); (e), HE (13); and (f), Ebo Van Weel website (14).
National Roads Policing Bureau (17) and Polizia Stradale (18) respectively operate in Ireland and Italy. Alongside managing traffic incidents, such TOs enforce highway laws, e.g., apprehend drink drivers and issue speeding fines.

In contrast, dedicated TO fleets, where we see the opportunity, are not as common as those formed from a police sub-unit; only two (HE and RWS) out of 16 highway agencies across Europe interviewed by Steenbruggen et al. (19) employ a dedicated TO fleet.

While on patrol, HE TOs, shown in Figure 1, either drive along designated patrol routes (typically the busier sections of the SRN) looking for incidents, or they remain stationary in waiting areas. TOs are instructed by operators in regional control centres (RCCs) to respond to incidents via a radio communications system installed in each vehicle. Although the fleet is not part of the UK emergency services, HE TOs have a number of extra powers, derived under the Traffic Management Act of 2004 (20), to quickly and safely clear incidents. For example, a HE TO may stop a vehicle, close highway lanes, and place temporary traffic signs. Furthermore, HE TOs are trained in first aid and CPR and can provide medical assistance at the roadside (21). A TO is often first to arrive at an incident and usually takes lead command, unless there is loss of life, when they fall under the command of the emergency services.

Similarly, the Weginspecteurs patrol and respond to traffic incidents across 5,200 km of highway in the Netherlands. Although the Weginspecteur fleet is separate from the police, some TOs in the Netherlands are enrolled as ‘Extraordinary Investigation Officers’ that may compile official reports against drivers for a number of driving offences, including ignoring red crosses on VMSs (indicating a lane closure) or driving on the hard shoulder. The Weginspecteur vehicles are fitted with a notable piece of equipment; a large vehicle-mounted variable message sign (VMS) to give direct instruction to highway users without relying on a roadside or gantry-mounted VMS.

**Vehicle Routing Problems**

Designing optimal vehicle routes is a widely studied operational research (OR) question known as the Vehicle Routing Problem (VRP). The VRP was first considered by Dantzig and Ramser (22) who developed a method to route a fleet of homogeneous fixed-capacity petrol trucks from a central depot to a series of delivery sites before returning back to the depot. Their method determines a set of shortest paths (one for each truck) along a transportation network defined as a graph of edges (road links) between demand nodes (delivery sites).

There are a number of VRP variations for specific transport logistic problems. One popular extension is the VRP with time windows (VRPTW) (23) in which nodes must be serviced within a given time window, to model deliveries to a store that must arrive within its opening hours, for example.

Standard VRPs assume that the demand is static and known before the vehicles begin their routes. However, in most real-world scenarios, the demand is dynamic (hail-and-ride taxi trip requests, for example) and vehicle routes are adjusted during operation (i.e., not at the depot); this extension is known as the dynamic VRP (DVRP).

In a recent review provided by Pillac et al. (24), the DVRP is categorised into two classes; deterministic and stochastic. In the first class of problem, the demand is revealed over time and an optimal solution may only be obtained for the current state. Typically, the deterministic demand problem is solved by periodically invoking a standard VRP (with the demand considered static at that time point) and correspondingly updating the vehicle routes.

The second class of problems consider demand with known statistics (computed from his-
torical data, for example). To solve this stochastic demand problem, an ensemble of hypothetical
future demand scenarios is sampled from the known demand distributions, and then solved as a set
of static VRPs. Routes that are averaged across the entire ensemble are then chosen.

When assigning a TO to an incident, a control centre operator considers the position of the
total fleet and dispatches a TO to minimise the incident response time, i.e., the time for the TO to
arrive at the incident. Such centralised vehicle dispatching systems (VDSs) are a specialisation of
the DVRP (with the additional constraint that demand should be serviced as soon as possible) and
are found in a number of demand-driven transportation systems such as hail-and-ride taxis (25)
and emergency repair vehicles (26).

Both a deterministic and stochastic VDS for a personal rapid transit (PRT) system are
considered by Lees-Miller and Wilson (27). The deterministic method assigns a vehicle as trips
are requested in sequence, whereas the stochastic method moves idle vehicles to stations to service
future (sampled) demand. On two test networks the stochastic method achieves up to a 96%
reduction in passenger waiting time compared against its deterministic counterpart.

The DVRP becomes complex for large transportation systems, such as highway networks,
and re-routing a vehicle multiple times throughout one journey may be unpractical and place a
heavy burden on the driver. Furthermore, standard DVRP-based approaches may not be appropri-
ate for vehicles with complicated operational constraints (with time, fuel, or mileage limits, for
example) such as a TO or police officer. Instead, complicated transportation problems are often
solved through simulations that may incorporate rich application-specific objectives.

The police patrol routing problem (PPRP), concerned with designing patrol routes to min-
imise the police’s response time to crime while maximally covering a patrol area to deter criminals,
is often considered via simulation. The PPRP has a strong analogy with our TO problem — both
aim to maximally cover an area (to deter criminals or capture asset data) while minimising response
time (to crime or traffic incidents).

The GAPatrol system developed by Reis et al. (28) is one simulation-based approach to the
PPRP. The authors developed a multi-agent simulation, based in the Fortaleza area of Brazil, in
which a group of criminals randomly try to commit crimes at predetermined targets while avoid-
ing a fleet of police officers. Their Genetic Algorithm-based method identifies a set of locations
for each police vehicle to visit in sequence that minimises the amount of crime committed in the
simulation. A similar simulation is presented by Melo et al. (29), however, in their work they con-
sider various ‘physical reorganising’ strategies; namely, how police officers drive between crime
hot spots. A short-routing strategy, where police officers drive along the shortest route between hot
spots, minimised the amount of crime.

Mobile Sensing for a Secondary Purpose and Maximal Coverage Systems
Mobile systems used for a secondary sensing purpose, such as the system proposed here, are less
widely deployed than those designed for a specific application, such as the specialised annual
survey vehicles. One example is the pothole patrol system (30) that uses data from accelerometers
installed in taxis across the city of Boston. The primary function of the taxis is not changed —
they pick up and drop off passengers as usual. However, by aggregating data collected from the
entire fleet, the system is able to automatically detect potholes across the city. Similarly, taxi fleets
have also been considered as city-wide VANET-based (vehicle ad-hoc network) communications
and traffic estimation systems (31, 32).

The objective of routing agents to maximally cover an area is another widely studied prob-
TABLE 1 (a) Three links in the HATRIS data set. The position of junctions M40 J12 and M40 J13 are recorded differently in rows 2 and 3 as the links represent different sides of the carriageway. (b) Six TO vehicles recorded in the TOFV data set. (c) Positions of three HE outstations, retrieved from UK government building database (36).

(a)

<table>
<thead>
<tr>
<th>Row</th>
<th>Description</th>
<th>Start junction</th>
<th>End junction</th>
<th>Start Easting [m]</th>
<th>Start Northing [m]</th>
<th>End Easting [m]</th>
<th>End Northing [m]</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>A38 between A513 and A5127</td>
<td>A38 A513</td>
<td>A38 A5127</td>
<td>417216</td>
<td>314346</td>
<td>414414</td>
<td>310549</td>
</tr>
<tr>
<td>2</td>
<td>M40 between M40 J12 and M40 J13</td>
<td>M40 J12</td>
<td>M40 J13</td>
<td>437249</td>
<td>254903</td>
<td>430537</td>
<td>260388</td>
</tr>
<tr>
<td>3</td>
<td>M40 between M40 J13 and M40 J12</td>
<td>M40 J13</td>
<td>M40 J12</td>
<td>430534</td>
<td>260412</td>
<td>437262</td>
<td>254918</td>
</tr>
</tbody>
</table>

(b)

<table>
<thead>
<tr>
<th>Vehicle Fleet Number</th>
<th>Make</th>
<th>Model</th>
<th>Outstation</th>
</tr>
</thead>
<tbody>
<tr>
<td>HE 1033</td>
<td>Land Rover</td>
<td>Discovery SDV6 3.0 DSE</td>
<td>Heston</td>
</tr>
<tr>
<td>HE 1034</td>
<td>Land Rover</td>
<td>Discovery SDV6 3.0 DSE</td>
<td>Heston</td>
</tr>
<tr>
<td>HE 0851</td>
<td>Mitsubishi</td>
<td>Shogun 3.2 DI-DC M</td>
<td>Heston</td>
</tr>
<tr>
<td>HE 0921</td>
<td>Mitsubishi</td>
<td>Shogun 3.2 DI-DC M</td>
<td>Samlesbury</td>
</tr>
<tr>
<td>HE 0945</td>
<td>Mitsubishi</td>
<td>Shogun 3.2 DI-DC M</td>
<td>Samlesbury</td>
</tr>
<tr>
<td>HE 0981</td>
<td>Land Rover</td>
<td>Discovery SDV6 3.0 DSE</td>
<td>Milnrow</td>
</tr>
</tbody>
</table>

(c)

<table>
<thead>
<tr>
<th>Outstation</th>
<th>Easting</th>
<th>Northing</th>
</tr>
</thead>
<tbody>
<tr>
<td>Knutsford</td>
<td>373433</td>
<td>378337</td>
</tr>
<tr>
<td>Milnrow</td>
<td>393036</td>
<td>411899</td>
</tr>
<tr>
<td>Watford</td>
<td>458602</td>
<td>270068</td>
</tr>
</tbody>
</table>

In the pioneering work of Machado et al. (33), the architectures and system considerations for maximal coverage multi-agent systems are tested within a simulation. A geographical area is de-constructed into a graph of nodes and edges, and the concept of node ‘idleness’ is introduced; that is, the time between successive node visits from an agent. Several subsequent works have continued to adopt the terminology of site or target idleness (34, 35). In (33) the agents try to minimise the total node idleness, and thus maximally cover the graph. Various routing strategies, including a greedy algorithm where each agent only considers their immediately neighbouring nodes, and more complex strategies where agents consider a larger neighbourhood of nodes are tested. Various communication styles are also tested, including decentralised methods where agents leave environmental ‘flags’ at the nodes for other agents to sense, centralised communication with a control centre, and communication via agent interaction. Our problem differs in that the TOs consider the idleness of the highway sections, rather than nodes. However, many of the topics covered are relevant considerations when deploying our TOs for TAM.

DATA SOURCES AND PREPARATION

We now focus our study on the HE TO fleet on the SRN in England, for which there is comprehensive data that we now describe. See Table 1.
Traffic Officers
Each vehicle in the HE TO fleet is recorded in the traffic officer fleet vehicles (TOFV) data set published by Highways England in 2018 (37). For each vehicle, the TOFV data set records its vehicle fleet identifier, make and model, and the outstation from which the vehicle begins and ends its patrol. In total, the TOFV consists of 234 vehicles that patrol from 32 outstations across the SRN.

Outstations
The name of each of the 32 TO outstations across the SRN is listed on the HE website (38). To determine the position of each outstation, their addresses were obtained via an online UK government building database (36) and converted to an Easting-Northing coordinate (39) to the nearest meter.

Strategic Road Network
The SRN is represented by the Highways Agency (HE’s name prior to 2014) traffic information service (HATRIS) links (40) that describe the junction-to-junction sections of the SRN. The links were initially used to report traffic flows between each junction until the service was upgraded in 2016 to provide live traffic information along shorter inter-junction highway sections (41).

For each link, its highway name (e.g., M6), and the start and end junction name (e.g., J1) and position, as an Easting-Northing coordinate to the nearest meter are recorded. Junctions along major highways are named by their junction number (e.g., M6 J1), however junctions along smaller highways are defined by the connecting road at the junction. For example, the A38 connects to the A515 (which is not part of the SRN) at junction A38 A515. In total, 2,510 links are recorded.

Graphical Highway Network Model
The recorded positions for each HATRIS junction with the same name are collected and their pairwise distances are computed. Those pairs for which the distance is greater than 4 km are then flagged for manual inspection and cleaning, since it seems likely that they are in fact different junctions, which happen to have the same name (for example, because pairs of ‘A’ roads might happen to meet at a number of different places).

An abstract graphical highway model (42) of the SRN is then constructed from the cleaned HATRIS links and junctions. The junctions provide the graph’s nodes and each HATRIS link represents a directed edge between those nodes. The driven distance for each edge is then modelled crudely by the straight-line distance between its nodes, based on their coordinates.

Unfortunately, the HATRIS data does not provide unambiguous junction coordinates: see Table 1(a), rows 2 and 3, which represent the opposing directions of the same dual carriageway section. Because of the lateral physical scale, different coordinates are recorded for the junction position, for each side of the carriageway. Our approach is thus simply to take the average of all of the coordinates provided for a given junction.

The resulting graph, see Figure 2, consists of 1,108 nodes and 2,376 directed edges, with a total length of 13,048 km. Highways England report a total network length of 13,760 km (43); the small difference can be accounted for by our neglect of link curvature.
FIGURE 2 (a) The SRN graph constructed from the HATRIS links and (b) two zoomed areas. For illustration purposes, each edge in this figure represents two directed edges (one for each direction of travel). The nodes closest to each outstation are depicted with orange squares.

CONCEPTS OF OPERATION

We now proceed to describe and simulate the proposed operational mode of the TOs. We will use the term FTO (future traffic officer) to distinguish between the new concepts of operation that we
Multi-lane highways usually have a central reservation and TOs can thus only re-route at junctions. Thus, we describe their operation in terms of a node-to-node routing methodology on the graphical model of the highway developed in the previous section.

Upon reaching a node, we suppose that each FTO will make one of the following decisions:

1. continue on the same highway;
2. join a new highway;
3. turn around and drive along the same highway in the opposite direction; or potentially,
4. stop for a rest period, albeit remaining available during this time for incident response.

The key idea is that different patrol regimes (that prioritise either incident response or asset management) will influence this routing decision.

In notation: the state of FTO $k$ is fully described by the node it is driving to, and the time it will arrive (and simultaneously depart from) there, respectively denoted by $n[k]$ and $a[k]$. The edge from node $i$ to $j$ has inter-visit time (idleness), $v(i, j)$, and driven distance, $d(i, j)$ (recall that this is modelled by the node-to-node straight-line distance). The distance matrix, $D$ describes the pairwise node shortest path length, and the trip time matrix is given by $T = D/s$, where we suppose that $s = 100$ kph is the constant FTO speed (i.e., the effect of congestion on travel times is not taken into account in this first model).

### Traffic Officer Patrol Regimes

We analyse three patrol regimes:

1. response (R1),
2. coverage (R2), or
3. random (R3).

The idea is that regime R1 aims to minimise the fleet’s collective incident response time, whereas regime R2 aims to maximise the fleet’s coverage for asset management. In regime R3, FTOs will select their next edge at random, and this thus acts as a base case which regimes R1 and R2 must out-perform.

Under regime R1, see Figure 3(a), FTO $k$ considers a neighborhood $N$ of close nodes, within a driven distance $\delta_n$ of node $n[k]$, and a set $F$ of close FTOs, within a driven distance $\delta_f$ of node $n[k]$. The idea is that FTOs continuously attempt to collectively minimise their drive time to any node within their near neighbourhood, in case an incident should occur there in the near future. This is achieved in a distributed way by each FTO selecting their next edge $e^*$ according to

$$e^* = \arg\min_e \sum_{n \in N} \tilde{r}_n,$$  

where

$$\tilde{r}_n = w_n \min \{ \{ r_e(k, n) \cap r(f, n) : f \in F \} \}$$

is the weighted minimum incident response time (WMIRT), and

$$r_e(k, n) = T(n[k], n_e) + T(n_e, n),$$

where $n_e$ denotes the destination node of edge $e$, is the response time of FTO $k$ to node $n$ if it
FIGURE 3 (a) Under regime R1 FTO 1 selects the edge to minimise the total weighted minimum response time to a nearby node neighbourhood (blue), from either itself, FTO 2, or FTO 3. (b) Under regime R2 FTO 1 selects the edge with the largest inter-visit time (red).

chooses to patrol along edge $e$. Note $w_n$ is a weight attached to node $n$ that models its relative frequency in experiencing incidents.

In contrast, in regime R2, see Figure 3(b), FTO $k$ selects the connecting edge with the largest inter-visit time, that is

$$ e^* = \arg \max_{e} [v(n[k], n_e)]. \quad (4) $$

In regime R3, the edge $e^*$ is chosen uniformly randomly from the allowed options. Once FTO $k$ has chosen which edge $e^*$ to patrol along, their state is updated according to $n[k] = n_{e^*}$ and $a[k] = a[k] + T(n[k], n_{e^*})$, and the inter-visit of the edge is updated such that $v(n[k], n_{e^*}) = 0$ minutes.

Incident Response Mechanism

From time to time, incidents will occur and the FTOs will break off from their usual patrolling pattern. We suppose that the Bell and Wong Nearest Neighbour (BWNN) heuristic (25) is employed to assign the nearest FTO to the incident and thus minimise the response time. In notation, given an incident at time $t$ along an edge from node $i$ to node $j$, BWNN assigns a single FTO, $k^*$, by

$$ k^* = \arg \min_k [r(k, i)]. \quad (5) $$
Here,

\[ r(k, i) = \max(0, a[k] - t) + T(n[k], i) + t_{inc}^i \]  

(6)

is the time for FTO \( k \) to drive to node \( i \), plus the time taken to drive from node \( i \) along the edge to the incident, denoted \( t_{inc}^i \). Correspondingly, \( t_{inc}^i \) denotes the time to drive from the incident to the edge’s destination node.

We suppose that asset monitoring continues while the selected FTO drives to the incident, and thus the inter-visit times of the edges along the shortest path from node \( n[k^\ast] \) to node \( i \), and the edge from node \( i \) to node \( j \), are set to zero as the FTO passes them. Moreover, we suppose that the assigned FTO is unavailable to respond to further incidents before clearing its current assignment. (Of course, one might consider much more sophisticated operations, where for example a multi-vehicle collision requires that multiple FTOs are deployed and taken from other less urgent assignments — this is beyond our current scope).

Once the incident is cleared, we assume that the FTO \( k^\ast \) drives to node \( j \), and thus \( n[k^\ast] = j \) and

\[ a[k^\ast] = t + r(k^\ast, i) + t_{TTC} + t_{inc}^j, \]  

(7)

and then resumes their usual patrolling regime. Here, \( t_{TTC} \) models the incident’s time-to-clear (TTC) (44).

**SIMULATION METHODOLOGY**

**Initialisation and Time Stepping**

At the beginning of the simulation, the inter-visit time of each edge is set to zero minutes; essentially, assets across the entire highway network begin in a perfect state (as if they have just been visited) and now require monitoring throughout the new patrol. Then, inspired by the day-night patrol pattern employed by HE and RWS (11, 45), each simulated FTO begins a 12-hour patrol shift at \( t = 0 \) minutes from the node closest to its outstation. Each FTO must return to this node by the end of the shift and is thus constrained by an operational radius that shrinks throughout the simulation. However, we allow an FTO to leave their operational area to respond to an incident.

While on patrol, HE TOs drive for half of their shift and rest at junctions or parking areas for the other half (HE, personal communication). To model this, our simulated FTOs repeatedly drive for one hour and then rest at a node for one hour. The first rest time of each FTO is randomly assigned within the first hour of the simulation to ensure that the entire FTO fleet is not resting at the same time. All simulated FTOs patrol under the same regime throughout the 12-hour patrol.

A simulation time step \( \Delta t = 1 \) minute is chosen and in each time step the update rules are applied to each simulated FTO in sequence. At time step \( m \), provided \( m\Delta t < a[k] \), FTO \( k \) will proceed along its current edge, provided it is not called to an incident. The node decision logic is then applied when \( m\Delta t \geq a[k] \), and the excess time, \( m\Delta t - a[k] \) is subtracted from the arrival time at its next node.

We experimented with varying values of \( \Delta t \), however the simulation was found to be robust (i.e., overall, the incident response times and edge inter-visit times were unchanged) when the time step was refined. A top-level simulation architecture is provided in Figure 4.
FIGURE 4 Simulation architecture. The FTOs patrol for a 12-hour shift and respond to incidents as they are generated (blue). Upon reaching a node, provided they are not resting, each FTO chooses which edge to patrol along next via the routing decision process (green) that is influenced by the FTO’s patrol regime.
Incident Model

Vehicle incidents on highways are often modelled by a Poisson process in space and time (46–48) that models each incident as an independent event. Furthermore, the rate at which a given edge suffers incidents is proportional to its length (48, 49).

In any time step, the probability of there being one event on an edge, from node $i$ to node $j$, is approximately $\lambda d(i, j)\Delta t$. Since this scales with small $\Delta t$, we take the usual approach and suppose that the probability of more than one event on a given edge in one time step is zero.

Each incident is randomly positioned along the edge and is given HE’s target TTC of $t_{\text{TTC}} = 60$ minutes (50).

In 2015, HE TOs attended 215,568 incidents, equal to 0.41 incidents per minute (51). Dividing this quantity by the total length of the SRN yields an incident rate $\lambda = 3.14 \times 10^{-5}$ incidents per km per minute.

Determination of Response Regime Parameters and Node Weightings

Under regime R1, each FTO considers the position of other nearby FTOs, to minimise the total WMIRT to a nearby node neighbourhood, see Equations 1 and 2. As the number of incidents on an edge is proportional to its driven distance, and each incident is randomly positioned along the edge, we model the weight $w_n$ of node $n$ by the total driven distance of its outgoing edges.

FTOs patrolling under regime R1 emerge into a kind-of formation where each FTO controls close to a response node; namely, a node from which the total WMIRT to all nodes in the neighbourhood is minimised.

When an FTO is assigned to an incident, they leave their response node and break the formation, then, the remaining unassigned FTOs emerge into a new formation. This reformation process is demonstrated in Figure 5 with a small fleet of three FTOs on a grid-like test graph, with 17 nodes and 28 edges each with length 10km. For this test case, $\delta_n$ is set such that each FTO’s node neighborhood is the entire graph, and $\delta_f$ is set such that each FTO may always communicate with the other two.

Three nodes in the test graph are given an increased weight of 100 whereas all other nodes have weight 10, and thus, the FTOs weight the minimum response time to these three nodes more heavily. Note that in this test case scenario, the weights are assigned independently of the edge length.

Intuitively, optimal R1 parameters will distribute the FTOs to response nodes across the SRN graph, so that an FTO may respond quickly to any incident. For the SRN graph considered in this paper, parameter values of $\delta_n = 80$km and $\delta_f = 80$km were found to distribute the FTOs (into desirable emergent formations) across the entire graph.

SIMULATION METRICS AND ANALYSIS

The incident response times and edge inter-visit times are compared over an ensemble of 50 simulations under regimes R1, R2, and R3 (150 in total). Four patrol routes from the ensemble are illustrated in Figure 6. Specifically, at simulation time $t$, we consider the ensemble-average incident response time (AIRT) and edge inter-visit time (AIVT), respectively computed by

$$AIRT(t) = \frac{1}{|N_{\text{inc}}|} \sum_{n \in N_{\text{inc}}} r(k^*, n)$$ (8)
FIGURE 5 Three FTOs on a test network under regime R1. Nodes $n_1, n_3$, and $n_{12}$ have weight 100, whereas the rest have weight 10 (see node legend and weights in the (c) panels). At the beginning of their patrol (a), the FTOs emerge into an initial formation. At $t = 40$ minutes the green FTO responds to an incident, and the two unassigned FTOs arrange themselves into a new formation (b).

\[ AIVT(t) = \frac{1}{|E|} \sum_{(i,j) \in E} v(i,j). \]  

Here, $N_{inc}$ denotes the start node of those edges where an incident occurs at time $t$ (and is attended
FIGURE 6 FTO patrol routes from the (a) Shepshed and (b) Milton Common outstations. The (i) and (ii) panels respectively correspond to patrols under regimes R1 and R2. Blue sections of highway indicate the FTO’s usual patrolling pattern whereas red indicates incident response: a darker colour illustrates where the FTO has regularly patrolled. All panels were generated via the OpenStreetMap mapping service (52).
Strain, Wilson and Littleworth

by FTO $k^*$, and nodes $i$ and $j$ respectively denote the start and end nodes of each edge in the set of SRN graph edges $\mathcal{E}$.

The AIVT is a suitable metric to consider for those assets that require immediate management (and thus continuous monitoring), such as debris on the highway surface. However, a daily asset data capture was specified by HE as a suitable monitoring frequency for assets whose condition might rapidly change but do not require immediate management, such as a broken street-light for example (HE, personal communication). Therefore, in addition to the AIVT, we also compute the ensemble-average percentage of edges that are visited at least once by an FTO (APV) during the 12-hour patrol. Specifically, at simulation time $t$, the APV is computed by

$$\text{APV}(t) = 100 \times \frac{1}{|\mathcal{E}|} \sum_{(i,j) \in \mathcal{E}} V(i,j),$$

where

$$V(i,j) = \begin{cases} 1 & \text{if } v(i,j) < t, \\ 0 & \text{otherwise}. \end{cases}$$

**Simulation Run-in and Return Period**

The ensemble AIVT and AIRT at each simulation time step are shown in Figure 7. The simulations exhibit a run-in period during the first 150 minutes of the patrol as the FTOs depart from their outstations and patrol on a small proportion of the SRN graph. The FTOs return to their outstations during a return period ($t > 650$ minutes). Consequently, under regime R2, the AIVT is increased as each FTO is restricted by their operational radius. On the other hand, under regime R1, the AIVT is decreased as the FTOs break their emergent formations. The AIRT is increased under all regimes during both the run-in and return period.

To assess the regimes while avoiding the impact of the run-in and return periods, we compute and compare the time-averaged AIRT (TAIRT) and AIVT (TAIVT) in the time-window $150 \leq t_w \leq 650$ minutes, such that,

$$\text{TAIRT} = \frac{1}{|t_w|} \sum_{t_w} \text{AIRT}(t_w),$$

and

$$\text{TAIVT} = \frac{1}{|t_w|} \sum_{t_w} \text{AIVT}(t_w).$$

Here, $|t_w|$ denotes the number of time steps in the time-window (500 for $\Delta t = 1$ minute). Similarly, we consider the APV at the end of the time-window (APVW),

$$\text{APVW} = \text{APV}(650 \text{ minutes});$$

essentially, this is the percentage of edges visited by at least one FTO before the return period.
FIGURE 7 (a) AIVT and (b) AIRT along with a 60 minute rolling-window average under each regime. The simulations exhibit a run-in period ($t < 150$ minutes) and a return period ($t > 650$ minutes) as the FTOs leave and return to their outstations.

Simulation Results
The TAIRT, TAIVT, and APVW under each regime are given in Table 2. Overall, regime R2 decreases the TAIVT by 106.46 minutes and the APVW by 15% compared with regime R1, whereas the TAIRT is only increased by 3.93 minutes. The ensemble distributions for the incident response time, and the edge inter-visit time at $t = 650$ minutes are provided in Figure 8.

The incident response mechanism (BWNN) is identical under all patrol regimes and thus the incident response time distributions have a similar profile for both regimes R1 and R2. In contrast, regime R1’s inter-visit time distribution exhibits a relatively heavy tail compared against regime R2, and 17% of edges are not visited at all during the patrol. In addition, an anomalously large number of inter-visit times from 0 – 10 minutes are exhibited under regime R1, because those edges connecting to response nodes are visited more regularly.
FIGURE 8 (a) Ensemble incident response times and (b) edge inter-visit time distributions at $t = 650$ minutes under regimes R1 and R2. The incident response mechanism (BWNN) is identical under both regimes and thus the response time distributions have a similar profile. Under regime R1, the FTOs frequently visit edges close to response nodes, thus, there are an anomalously large number of inter-visit times from 0 – 10 minutes, and a large proportion of edges that are not visited at all.

Varying Fleet Sizes

Figure 9 shows the TAIRT, TAIVT, and APVW achieved by each regime for varying fleet sizes (averaged across 30 simulations per fleet size). When the fleet size is small, the number of incidents overwhelms the fleet and a queue of incidents for each FTO to attend builds. Consequently, the FTO patrol is identical under all regimes; each FTO attends and clears an incident, and then attends the next and so on. As a result, all regimes achieve an asymptotically large TAIRT, and small TAIVT and APVW.

At the other extreme, for large fleet sizes, the number of FTOs approximates the number of nodes in the graph (1,108). In this case, the fleet is equally well distributed under either regime and the metrics tend to an asymptotic value. The key result shown in Table 2 is reproduced for all realistic fleet sizes — the FTO fleet can be deployed to efficiently collect asset data across the SRN, while only increasing the fleet’s incident response time by a few minutes.
TABLE 2 TAIRT, TAIVT, and APVW achieved by an FTO fleet of 234 vehicles. The difference row shows the difference between regimes R2 and R1. Overall, regime R2 reduces the TAIVT by 106.46 minutes and increases the APV by 15%, whereas, the TAIRT is only increased by 3.93 minutes.

<table>
<thead>
<tr>
<th>Regime</th>
<th>TAIRT [minutes]</th>
<th>TAIVT [minutes]</th>
<th>APVW [%]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Response (R1)</td>
<td>8.52</td>
<td>186.74</td>
<td>83</td>
</tr>
<tr>
<td>Coverage (R2)</td>
<td>12.45</td>
<td>80.28</td>
<td>98</td>
</tr>
<tr>
<td>Difference</td>
<td>3.93</td>
<td>-106.46</td>
<td>15</td>
</tr>
<tr>
<td>Random (R3)</td>
<td>13.41</td>
<td>135.16</td>
<td>93</td>
</tr>
</tbody>
</table>

DISCUSSION

1. The TAIRT achieved by regime R2 is only a 0.96 minute improvement on regime R3 (random) – this result is unsurprising as neither regime tries to position the fleet for fast incident response. However, regime R2 still achieves a TAIRT well under HE’s current average incident response time of 17 minutes, and only 2.45 minutes over their target response time of 10 minutes (53).

2. Regime R1 reduces the TAIRT by 3.93 minutes (compared against regime R2) but only minimises the inter-visit times of edges close to response nodes or along the shortest path to an incident. Consequently, the TAIVT is significantly larger under regime R1, compared with regimes R2 and R3. Furthermore, on average, 17% of edges are not visited at all under regime R1, whereas only 2% of edges are not visited under regime R2. The key result, that regime R2 significantly reduces the TAIVT and APVW but only increases the TAIRT by a few minutes, is reproduced for all realistic fleet sizes (see Figure 9), and thus, our simulations show the potential to use TO fleets for asset management.

3. In this work, we have considered a base-case regime where each FTO chooses an edge at random (R3). Of course, there are other regimes that may result in longer incident response times and reduced coverage — every FTO remaining at their outstation, for example. However, regime R3 employs the same node-to-node routing methodology as that in regimes R1 and R2, and thus seems an appropriate base case.

4. The simulation has a number of simplifications, and future refinements should aim to further reflect the behavioural details of real-world TO patrols. For example, the simulated FTOs drive with a constant speed and we do not consider the effect of traffic congestion. Therefore, historical traffic data might be incorporated to reflect realistic driving conditions. In addition, the FTOs presently make an instantaneous routing decision at a node without changing their speed. Although this is reasonably realistic for TOs who choose to remain on the same highway, TOs who merge onto a different highway may slow down on a slip road or stop at a traffic lighted junction and thus incur a delay.

5. Presently, we model the number of incidents on an edge as being proportional to its driven distance, with a constant of proportionality that is common to the entire network. However, an improved incident model might be calibrated with historical data concerning the number of incidents attended by TOs on each highway section. While public road safety data sets exist that record the number of vehicle crashes and collisions across the SRN (54), TOs respond to a number of other types of incidents (breakdowns and removing abandoned vehicles, for example). Therefore, efforts should be made to liaise with highway agencies to determine the number of incidents attended by
FIGURE 9 Panels (a) to (c) respectively show the TAIRT, TAIVT, and APVW for varying fleet sizes. The (i) panels show each metric across a wide range of fleet sizes whereas the (ii) panels are cropped to show more realistic fleet sizes.

1 TOs at a section-by-section level.
2 Our proposed concept of operation assumes that the SRN and FTO vehicles are equipped with the necessary communications infrastructure for the FTOs to patrol under each regime. When
an FTO makes a routing decision (under either regime R1 or R2) they require real-time data; regime R2 requires the edge inter-visit times, and regime R1 requires knowledge of the nearby FTOs and nodes around each FTO. Through dialogue with HE, it is understood that each RCC already captures live telemetry data for HE TO vehicles via an installed GPS-enabled device (HE, personal communication). Therefore, the data required by either regime can be computed in real-time (the inter-visit times would be updated as a TO drives along a highway section) with the current technology.

To implement an operational system, a highway agency must decide how to compute and communicate the navigation instructions to each TO. This may be achieved by one of two ways:

1. the routing decision for each TO is computed in the RCC and then communicated back to the TO, either by radio or a connected (vehicle to RCC) navigation system; or,
2. the required data is sent back to the TO vehicle and a routing decision is computed by a small device installed in the vehicle.

There are a number of other possible operational constraints on the TO patrols that are not considered in our simulations, such as maximum mileage constraints, for example. In addition, the FTOs in our simulations are all assumed to patrol under the same regime. In reality, TOs may have varying roles; for example, only some of the RWS Weginspecteurs may compile official incident reports. Therefore, future work might incorporate non-homogeneous FTO fleets, using different patrol regimes.

CONCLUSION

In this paper, patrol routing strategies are proposed and simulated to consider whether a fleet of traffic officers (TOs) may be used for TAM alongside their primary role of incident management. A case study of the Highways England (HE) TO fleet on the strategic road network (SRN) in England, UK, is considered to investigate, as a proof-of-concept, whether data from dash cams installed in TO vehicles might be used to frequently capture asset data across an entire highway network.

Junction-to-junction links in the HATRIS data set were used to build a graphical model of the SRN. A simulator was built to deploy the HE TO fleet who respond to dynamically generated incidents. Within the simulation, TOs patrol under one of three regimes that influences their routing decisions at each node (junction). Three regimes were considered: response, coverage, and random. The first aims to minimise the fleet’s incident response time and acts as a best case to compare the other regimes’ incident response times against. The second aims to maximise the fleet’s coverage for asset management. The third regime, where TOs make random routing decisions, acts as a base case.

To determine the feasibility of the proposed patrol routes, the incident response times and the edge inter-visit times were computed. Overall, our simulations showed that the TOs (with varying fleet sizes) can monitor assets across the SRN while only increasing the incident response time by a few minutes.

We have demonstrated that TOs can be used for frequent asset data capture, alongside their primary role of incident response. To employ our proposed patrol routing strategies in operation, the increase in incident response time must be weighed up by decision makers, who also need to consider the costs of the required sensor and communications infrastructure. However, note that for the HE TO fleet considered in this paper, the required sensing capability (the dash cams) and much of the communication infrastructure (radio and telemetry data sent to a regional control...
centre) is already in place.

ACKNOWLEDGEMENTS
The authors would like to acknowledge the support provided by Jacobs, Highways England, and the Engineering and Physical Sciences Research Council (EPSRC) that have made the work presented in this paper possible. This work was supported by a National Productivity Investment Fund (NPIF) (award number 2107418) with contributions from Jacobs.

AUTHOR CONTRIBUTIONS
The authors confirm contributions to the paper as follows: study conception and design: T.S., R.E.W., R.L.; data collection T.S., R.E.W., R.L.; analysis and interpretation of results: T.S., R.E.W., R.L.; draft manuscript preparation: T.S., R.E.W., R.L. All authors reviewed the results and approved the final version of the manuscript.
REFERENCES


